

of 11 non-subtype B HIV-1 viruses isolated from primary isolates collected at major epicenters of the global AIDS pandemic. The nucleotide sequences of these 11 viruses are shown in Figures 13A-13Z (SEQ ID NOS: 1 to 11).

Please replace the paragraph beginning at page 14, line 26, with the following rewritten paragraph:

F2  
The present invention relates to nucleic acids having the genomic sequence of any one of the 11 molecular clones for non-subtype HIV-1 isolates of this invention as shown in Figures 13A-13Z (SEQ ID NOS: 1 to 11), as well as fragments (or partial sequences) thereof. The invention also relates to nucleic acids having complementary (or antisense) sequences to the sequences shown in Figures 13A-13Z (SEQ ID NOS: 1 to 11), as well as fragments (or partial sequences) thereof. Partial sequences may be obtained by various methods, including restriction digestion of nucleic acids with sequences shown in Figures 13A-13Z (SEQ ID NOS: 1 to 11), PCR amplification, and direct synthesis. Partial sequences may be all or part of the LTR and/or other untranslated regions of the genomes of one or more of the 11 viral clones of this invention, and/or all or part of the genes encoding the Gag, Pol, Vif, Vpr, Env, Tat, Rev, Nef and Vpu proteins and/or complementary (or antisense) sequences thereof. Nucleic acids of the invention also include cDNA, mRNA, and other nucleic acids derived from the genomic sequences of one or more of these 11 HIV-1 clones. Sequences of the genes encoding Gag, Pol, Vif, Vpr, Env, Tat, Rev, Nef and Vpu are identified in Figures 13A-13Z.

Please replace the paragraph beginning at page 16, line 18, with the following rewritten paragraph:

F3  
The nucleic acid probes used in the detection methods set forth above are derived from nucleic acid sequences shown in Figures 13A-13Z (SEQ ID NOS: 1 to 11). The size of such probes is at least 10-12 bases long, more usually at least about 19 bases long, more usually from about 200 to about 500 bases, and often exceeding about 1000 bases.

Please replace the paragraph beginning at page 17, line 7, with the following rewritten paragraph:

F4  
The nucleic acid probes used in the detection methods set forth above are derived from sequences substantially homologous to one or more of the sequences shown in Figures 13A-13Z (SEQ ID NOS: 1 to 11), or their complementary sequences. By "substantially homologous", as used throughout the specification and claims to describe the nucleic acid sequence of the present invention, is meant a high level of homology between the nucleic acid sequence and one or more of the sequences of Figures 13A-13Z (SEQ ID NOS: 1 to 11), or its complementary sequence. Preferably, the level of homology is in excess of 80%, more preferably in excess of 90%, with a preferred nucleic acid sequence being in excess of 95% homologous with a portion of one or more of the sequences shown in Figures 13A-13Z (SEQ ID NOS: 1 to 11), or its complement. The size of such probes is usually at least 20 nucleotides, more usually from about 200 to 500 nucleotides, and often exceeding 1000 nucleotides.

Please replace the paragraph beginning at page 17, line 28, with the following rewritten paragraph:

The methods for analyzing the RNA for the presence of the viruses of this invention include Northern blotting (94), dot and slot hybridization, filter hybridization (95), Rnase protection (93), and reverse-transcription polymerase chain reaction (RT-PCR)(96). A preferred method is RT-PCR. In this method, the RNA can be reverse transcribed to first strand cDNA using a nucleic acid primer or primers derived from one or more of the nucleotide sequences shown in Figures 13A-13Z (SEQ ID NOS: 1 to 11). Once the cDNAs are synthesized, PCR amplification is carried out using pairs of primers designed to hybridize with sequences in the genomes of one or more of the non-subtype B HIV-1 viruses of this invention which are an appropriate distance apart (at least about 50 bases) to permit amplification of the cDNA and subsequent detection of the amplification product. Each primer of a pair is a single-stranded nucleic acid of about 20 to about 60 bases in length where one primer (the "upstream" primer) is complementary to the original RNA and the second primer (the "downstream" primer) is complementary to the first strand of cDNA generated by reverse transcriptions of the RNA. The target sequence is generally about 100 to about 300 bases in length but can be as large as 500-1500 bases or more, e.g., 9,000 bases. Optimization of the amplification reaction to obtain sufficiently specific hybridization to the nucleotide sequences of these viruses is well within the skill in the art and is preferably achieved by adjusting the annealing temperature.

Please replace the paragraph beginning at page 21, line 27, with the following rewritten paragraph:

~~The polypeptides of this invention consist of at least 6-12 amino acids, more preferably at least 3-18 amino acids, even more preferably at least 19-24 amino acids and most preferably at least 25-30 amino acids encoded by or otherwise derived from, any one of the genomic sequences shown in Figures 13A-13Z (SEQ ID NOS: 1 to 11).~~

Please replace the paragraph beginning at page 31, line 9, with the following rewritten paragraph:

~~The present invention further relates to computer-generated alignments of any or more of the nucleotide sequences shown in Figures 13A-13Z (SEQ ID NOS: 1 to 11). Computer analysis of the nucleotide sequences, such as the one shown in Figure 13A-13Z, can be carried out using commercially available computer program known to one skilled in the art.~~

Please replace the paragraph beginning at page 31, line 14, with the following rewritten paragraph:

~~In one embodiment, the sequences shown in Figures 13A-13Z (SEQ ID NOS: 1 to 11) are aligned by the computer program CLUSTAL (67) and adjusted with multiple-aligned sequence editor (12). The computer analysis results in the distribution of 11 sequences into various genotypes. Five of these sequences represent non-recombinant members of HIV-1 subtypes, and the other six sequences represent HIV-1 intersubtype recombinants.~~

Please replace the paragraph beginning at page 33, line 11, with the following rewritten paragraph:

(Near) full length HIV-1 genomes were amplified from short-term cultured PMBC DNA essentially as described (18,56) using the GeneAmp XL kit (Perkin Elmer Cetus, Foster City, Calif.) and primers spanning the tRNA primer binding site (upstream primer UPIA: 5'-AGTGGCGCCCCGAACAGG-3') (SEQ ID NO: 109) and the R/U5 junction in the 3' long terminal repeat (downstream primer Low2: 5'-TGAGGCTTAAGCAGTGGGTTTC-3') (SEQ ID NO: 110). Some isolates were amplified with primers containing MluI restriction enzyme sites to facilitate subsequent subcloning into plasmid vectors (upstream primer UPIAMluI: 5'-TCTCTacgcgtGGCGCCCGAACAGGGAC-3' (SEQ ID NO: 111); downstream primer Low1MluI: 5'-ACCAgacgcgtACAACAGACGGGCACACACTA-CTT-3' (SEQ ID NO: 112); lower case letters indicate the MluI restriction site). Whenever possible, PBMC DNAs were diluted prior to PCR analysis to attempt amplification from single proviral templates. Cycling conditions included a hot start (94°C, 2 min), followed by 20 cycles of denaturation (94°C; 30 sec) and extension (68°C; 10 min), followed by 17 cycles of denaturation (94°C; 30 sec) and extension (68°C, 10 min) with 15 second increments per cycle. PCT products were visualized by agarose gel electrophoresis and subclone into pCRII by T/A overhang or following cleavage with MluI into a modified pTZ18 vector (pTZ18MluI) containing a unique MluI site in its polylinker. Transformations were performed in INVαF<sup>+</sup> cells, and colonies were screened by restriction enzyme digestion for full length inserts (transformation efficiencies were generally poor, yielding only a few recombinant colonies, however, once subcloned, full length genomes

were stable in their respective vectors). One full length clone per isolate was randomly chosen for subsequent sequence analysis.

Please replace the paragraph beginning at page 50, line 22, with the following rewritten paragraph:

Using primers homologous to the tRNA primer binding site (5'-TCTCT-acgcgtGGCGCCCGAACAGGGAC-3' (SEQ ID NO: 111), lower case letters indicate an MluI site) and the polyadenylation signal in the 3' LTR (5'-ACCAGacgcgtACAACAGACGGG-CACACACTACTT-3')(SEQ ID NO: 112), long range PCR was used to amplify near full length genomic fragments, which contained all coding and regulatory regions except for 102 bp of 5' unique LTR sequences (U5)(for methodological details concerning the long range PCR approach see refs. 18, 56, 79). Amplification products were subcloned into a plasmid vector, mapped by restriction enzyme digestion, and one clone (94CY032.3) was selected for further analysis. A 694 bp fragment spanning the remainder of the LTR was amplified separately using a semi-nested approach (18).